

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANTS: Seed, Brian et al.

(ii) TITLE OF INVENTION: Targeted Cytolysis of HIV-Infected Cells by Chimeric CD4 Receptor-Bearing Cells

(iii) NUMBER OF SEQUENCES: 27

(iv) CORRESPONDENCE ADDRESS:

(A) ADDRESSEE: Fish & Richardson
(B) STREET: 225 Franklin Street
(C) CITY: Boston
(D) STATE: MA
(E) COUNTRY: USA
(F) ZIP: 02110-2804

(v) COMPUTER READABLE FORM:

(A) MEDIUM TYPE: 3.5" Diskette, 1.44 Mb
(B) COMPUTER: IBM PS/2 Model 50Z or 55SX
(C) OPERATING SYSTEM: IBM P.C. DOS (Version 3.30)
(D) SOFTWARE: Wordperfect (Version 5.0)

(vi) CURRENT APPLICATION DATA:

(A) APPLICATION NUMBER:
(B) FILING DATE:
(C) CLASSIFICATION:

(vii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 07/847,566
(B) FILING DATE: March 6, 1992
(C) CLASSIFICATION:

(viii) PRIOR APPLICATION DATA:

(A) APPLICATION NUMBER: 07/665,961
(B) FILING DATE: March 7, 1991
(C) CLASSIFICATION:

(ix) ATTORNEY/AGENT INFORMATION:

(A) NAME: Clark, Paul T.
(B) REGISTRATION NUMBER: 30,162
(C) REFERENCE/DOCKET NUMBER: 00786/212001

(x) TELECOMMUNICATION INFORMATION:

(A) TELEPHONE: (617) 542-5070
(B) TELEFAX: (617) 542-8906
(C) TELEXI: 200154

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 1728 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

ATGAACCGGG GAGTCCTTT TAGGCACITG CTTCTGGTGC TGCAACTGGC 50
GCTCCTCCC GCAGCCACTC AGGGAAACAA AGTGGTGCTG GGCAAAAAAG 100
GGGATACAGT GGAACGTGACC TGTACAGCTT CCCAGAAGAA GAGCATAACAA 150
TTCCACTGGA AAAACTCCAA CCAGATAAAG ATTCTGGAA ATCAGGGCTC 200
CTTCTTAAC TAAAGGTCCAT CCAAGCTGAA TGATCGCGCT GACTCAAGAA 250
GAAGCCCTTG GGACCAAGGA AACTCCCCC TGATCATCAA GAATCTTAAG 300
ATAGAAGACT CAGATACTTA CATCTGTGAA GTGGAGGACC AGAAGGAGGA 350
GGTGCAATTG CTAGTGTTCG GATTGACTGC CAACTCTGAC ACCCACCTGC 400
TTCAGGGGCA GAGCCTGACC CTGACCTTGG AGAGCCCCC TGGTAGTAGC 450
CCCTCAGTGC AATGTAGGAG TCCAAGGGGT AAAAACATAC AGGGGGGGAA 500
GACCCTCTCC GTGTCTCAGC TGGAGCTCCA GGATAGTGGC ACCTGGACAT 550
GCACTGTCTT GCAGAACCAAG AAGAAGGTGG AGTTCAAAAT AGACATCGTC 600
GTGCTAGCTT TCCAGAAGGC CTCCAGCATA GTCTATAAGA AAGAGGGGGAA 650
ACAGGTGGAG TTCTCCTTCC CACTCGCCTT TACAGTTGAA AAGCTGACGG 700
GCAGTGGCGA GCTGTGGTGG CAGGGGGAGA GGGCTTCCTC CTCCAAGTCT 750
TGGATCACCT TTGACCTGAA GAACAAGGAA GTGTCTGTAA AACGGGTTAC 800
CCAGGACCT AAGCTCCAGA TGGGCAAGAA GCTCCCGCTC CACCTCACCC 850
TGGCCCCAGGC CTTGCCTCAG TATGCTGGCT CTGGAAACCT CACCTGGCC 900
CTTGAAGCGA AAACAGGAAA GTTGCATCAG GAAGTGAACC TGGTGGTGAT 950
GAGAGCCACT CAGCTCCAGA AAAATTGAC CTGTGAGGTG TGGGGACCCA 1000
CCTCCCCCTAA GCTGATGCTG AGCTTGAAAC TGGAGAACAA GGAGGCAAAG 1050
GTCTCGAACG GGGAGAAGCC GGTGTGGGTG CTGAACCCCTG AGGGGGGGAT 1100
GTGGCAGTGT CTGCTGAGTG ACTCGGGACA GGTCTGCTG GAATCCAACA 1150
TCAAGGTTCT GCCCACATGG TCCACCCGG TGCACGCGGA TCCCAAACTC 1200
TGCTACTTGC TAGATGGAAT CCTCTTCATC TACGGAGTCA TCATCACAGC 1250
CCTGTACCTG AGAGCAAAAT TCAGCAGGAG TGCAGAGACT GCTGCCAAC 1300
TGCAGGACCC CAACCAAGCTC TACAATGAGC TCAATCTAGG GCGAAGAGAG 1350
GAATATGACG TCTTGGAGAA GAAGGGGGCT CGGGATCCAG AGATGGGAGG 1400
CAAACAGCAG AGGAGGAGGA ACCCCCCAGGA AGGCGTATAAC AATGCACTGC 1450
AGAAAGACAA GATGCCAGAA GCCTACAGTG AGATCGGCAC AAAAGGGCGAG 1500

AGGCAGGAGAG GCAAGGGGCA CGATGCCCTT TACCAAGGACA GCCACCTCCA	1550
AGCAGTGCAG TTCGGGAACA GAAGAGAGAG AGAAGGTTCA GAACTCACAA	1600
GGACCCCTGG GTTAAGAGCC CGCCCCAAAG GTGAAAGCAC CCAGCAGAGT	1650
AGCCAATCCT GTGCCAGCGT CTTCAGCATE CCCACTCTGT GGAGTCCATG	1700
GCCACCCAGT AGCAGCTCCC AGCTCTAA	1728

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1389 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

ATGAACCGGG GAGTCCCTTT TAGGCACCTTG CTTCTGGTGC TCCAACCTGGC	50
GCTCCTCCCA GCAGCCACTC AGGGAAACAA AGTGGTGCTG GGCAAAAAAG	100
GGGATACAGT GGAACGTGACC TGTACAGCTT CCCAGAAGAA GAGCATAACAA	150
TTCCACTGGA AAAACTCCAA CCAGATAAAG ATTCTGGAA ATCAGGGCTC	200
CTTCTTAACT AAAGGTCCAT CCAAGCTGAA TGATCGCGCT GACTCAAGAA	250
GAAGCCTTTG GGACCAAGGA AACTCCCCC TGATCATCAA GAATCTTAAG	300
ATAGAAGACT CAGATACTTA CATCTGTGAA GTGGAGGACC AGAAGGAGGA	350
GGTGCAATTG CTAGTGTTCG GATTGACTGC CAACTCTGAC ACCCACCTGC	400
TTCAGGGGCA GAGCCTGACC CTGACCTTGG AGAGCCCCC TGGTAGTAGC	450
CCCTCAGTGC AATGTAGGAG TCCAAGGGT AAAAACATAC AGGGGGGGAA	500
GACCCTCTCC GTGTCTCAGC TGGAGCTCCA GGATAGTGGC ACCTGGACAT	550
GCACTGTCTT GCAGAACCCAG AAGAAGGTGG AGTTCAAAAT AGACATCGTG	600
GTGCTAGCTT TCCAGAAGGC CTCCAGCATA GTCTATAAGA AAGAGGGGGAA	650
ACAGGTGGAG TTCTCCCTCC CACTCGCCTT TACAGTTGAA AAGCTGACGG	700
GCAGTGGCGA GCTGTGGTGG CAGGGGAGA GGGCTTCCTC CTCCAAGTCT	750
TGGATCACCT TTGACCTGAA GAACAAGGAA GTGTCTGTAA AACGGGTTAC	800
CCAGGACCT AAGCTCCAGA TGGGCAAGAA GCTCCCGCTC CACCTCACCC	850
TGCCCCAGGC CTTCAGCAG TATGCTGGCT CTGGAAACCT CACCCCTGGCC	900
CTTGAAGCGA AAACAGGAAA GTTGCATCAG GAAGTGAACC TGGTGGTGAT	950

GAGAGCCACT CAGCTCCAGA AAAATTGAC CTGTGAGGTG TGGGGACCCA	1000
CCTCCCCCTAA GCTGATGCTG AGCTTGAAC TGGAGAACAA GGAGGCAAAG	1050
GTCTCGAACG CGGAGAAGCC GGTGTGGTG CTGAACCTG AGGCCGGGAT	1100
GTGGCAGTGT CTGCTGAGTG ACTCGGGACA GGTCTGCTG GAATCCAACA	1150
TCAAGGTTCT GCCCACATGG TCCACCCCGG TGCACGCGGA TCCGCAGCTC	1200
TGCTATATCC TGGATGCCAT CCTGTTTTG TATGGTATTG TCCTTACCCCT	1250
GCTCTACTGT CGACTCAAGA TCCAGGTCCG AAAGGCAGAC ATAGCCAGCC	1300
GTGAGAAATC AGATGCTGTC TACACGGGC TGAACACCCG GAACCAGGAG	1350
ACATATGAGA CTCTGAAACA TGAGAAACCA CCCAATAG	1389

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 1599 base pairs
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: double
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

ATGAACCGGG GAGTCCCTTT TAGGCACTTG CTTCTGGTGC TGCAACTGGC	50
GCTCCTCCCA GCAGCCACTC AGGGAAACAA AGTGGTGCTG GGCRAAAAAG	100
GGGATACAGT GGAAC TGACC TGTACAGCTT CCCAGAAGAA GAGCATAACAA	150
TTCCACTGGA AAAACTCCAA CCAGATAAAG ATTCTGGAA ATCAGGGCTC	200
CTTCTTAACAAAGGTCCAT CCAAGCTGAA TGATCGCGCT GACTCAAGAA	250
GAAGCCTTTG GGACCAAGGA AACTCCCCC TGATCATCAA GAATCTTAAG	300
ATAGAAAGACT CAGATACTTA CATCTGTGAA GTGGAGGACC AGAAGGAGGA	350
GGTGCAATTG CTAGTGTTCG GATTGACTGC CAACTCTGAC ACCCACCTGC	400
TTCAGGGGCA GAGCCTGACC CTGACCTTGG AGAGCCCCC TGTTAGTAGC	450
CCCTCAGTGC AATGTAGGAG TCCAAGGGGT AAAAACATAC AGGGGGGGAA	500
GACCCCTCTCC GTGTCTCAGC TGGAGCTCCA GGATAGTGGC ACCTGGACAT	550
GCACGTGCTT GCAGAACCGAG AAGAAGGTGG AGTTCAAAAT AGACATCGTG	600
GTGCTAGCTT TCCAGAAGGC CTCCAGCATA GTCTATAAGA AAGAGGGGGA	650
ACAGGTGGAG TTCTCCCTCC CACTCGCCTT TACAGTTGAA AAGCTGACGG	700
GCAGTGGCGA CCTGTGGTGG CAGGGGGAGA GGGCTTCCTC CTCCAAGTCT	750

TGGATCACCT TTGACCTGAA GAAACAGGAA GTGTCTGTAA AACGGGTTAC	800
CCAGGACCTT AAGCTCCAGA TGGGCJAGAA GCTCCCGCTC CACCTCACCC	850
TGCCCCAGGC CTTGCCTCAG TATGCTGGCT CTGGAAACCT CACCCCTGGCC	900
CTTGAAGCGA AAACAGGAAA GTTGCATCAG GAAGTGAACC TGGTGGTGAT	950
GAGAGCCACT CAGCTCCAGA AAAATTGAC CTGTGAGGTG TGGGACCCA	1000
CCTCCCCCTAA GCTGATGCTG AGCTTGAAAC TGGAGAACAA GGAGGCAAAG	1050
GTCTCGAACG CCGAGAAGCC GGTGTGGTG CTGAACCTG AGGCGGGAT	1100
GTGGCAGTGT CTGCTGAGTG ACTCGGGACA GGTCCCTGCTG GAATCCAACA	1150
TCAAGGTTCT GCCCACATGG TCCACCCCCG TGCACGGCGA TCCCACACTC	1200
TGCTACCTGC TGGATGGAAT CCTCTTCATC TATGGTGTCA TTCTCACTGC	1250
CTTGTTCCTG AGAGTGAAGT TCAGCAGGAG CGCAGAGCCC CCCGGCTTAC	1300
AGCAGGGCCA GAAACAGCTC TATAACGAGC TCAATCTAGG ACGAAGAGAG	1350
GAGTACGATG TTTTGGACAA GAGACGTGGC CGGGACCCCTG AGATGGGGGG	1400
AAAGCCGAGA AGGAAGAACCC CTCAGGAAGG CCTGTACAAT GAACTGCAGA	1450
AAGATAAGAT CGCGGAGGCC TACAGTGAGA TTGGGATGAA AGGGAGGCC	1500
CGGAGGGGCC AGGGGCACGA TGGCCTTAC CAGGGTCTCA GTACAGGCCAC	1550
CAAGGACACC TACGACGCC TTCACATGCA GGCCCTGCC CCGCTAA	1599

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 575 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

Met	Asn	Arg	Gly	Val	Pro	Phe	Arg	His	Leu	Leu	Leu	Val	Leu	Gln	Leu
1					5				10				15		
Ala	Leu	Leu	Pro	Ala	Ala	Thr	Gln	Gly	Asn	Lys	Val	Val	Leu	Gly	Lys
					20				25				30		
Lys	Gly	Asp	Thr	Val	Glu	Leu	Thr	Cys	Thr	Ala	Ser	Gln	Lys	Lys	Ser
					35			40				45			
Ile	Gln	Phe	His	Trp	Lys	Asn	Ser	Asn	Gln	Ile	Lys	Ile	Leu	Gly	Asn
					50			55				60			
Gln	Gly	Ser	Phe	Leu	Thr	Lys	Gly	Pro	Ser	Lys	Leu	Asn	Asp	Arg	Ala
					65			70				75			80
Asp	Ser	Arg	Arg	Ser	Leu	Trp	Asp	Gln	Gly	Asn	Phe	Pro	Leu	Ile	Ile
					85			90				95			
Lys	Asn	Leu	Lys	Ile	Glu	Asp	Ser	Asp	Thr	Tyr	Ile	Cys	Glu	Val	Glu
					100			105				110			
Asp	Gln	Lys	Glu	Glu	Val	Gln	Leu	Leu	Val	Phe	Gly	Leu	Thr	Ala	Asn
					115			120				125			

Ser Asp Thr His Leu Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu Glu
130 135 140
Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly
145 150 155 160
Lys Asn Ile Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu
165 170 175
Gln Asp Ser Gly Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys
180 185 190
Val Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser
195 200 205
Ser Ile Val Tyr Lys Lys Glu Gly Glu Gln Val Glu Phe Ser Phe Pro
210 215 220
Leu Ala Phe Thr Val Glu Lys Leu Thr Gly Ser Gly Glu Leu Trp Trp
225 230 235 240
Gln Ala Glu Arg Ala Ser Ser Lys Ser Trp Ile Thr Phe Asp Leu
245 250 255
Lys Asn Lys Glu Val Ser Val Lys Arg Val Thr Gln Asp Pro Lys Leu
260 265 270
Gln Met Gly Lys Lys Leu Pro Leu His Leu Thr Leu Pro Gln Ala Leu
275 280 285
Pro Gln Tyr Ala Gly Ser Gly Asn Leu Thr Leu Ala Leu Glu Ala Lys
290 295 300
Thr Gly Lys Leu His Gln Glu Val Asn Leu Val Val Met Arg Ala Thr
305 310 315 320
Gln Leu Gln Lys Asn Leu Thr Cys Glu Val Trp Gly Pro Thr Ser Pro
325 330 335
Lys Leu Met Leu Ser Leu Lys Leu Glu Asn Lys Glu Ala Lys Val Ser
340 345 350
Lys Arg Glu Lys Pro Val Trp Val Leu Asn Pro Glu Ala Gly Met Trp
355 360 365
Gln Cys Leu Leu Ser Asp Ser Gly Gln Val Leu Leu Glu Ser Asn Ile
370 375 380
Lys Val Leu Pro Thr Trp Ser Thr Pro Val His Ala Asp Pro Lys Leu
385 390 395 400
Cys Tyr Leu Leu Asp Gly Ile Leu Phe Ile Tyr Gly Val Ile Ile Thr
405 410 415
Ala Leu Tyr Leu Arg Ala Lys Phe Ser Arg Ser Ala Glu Thr Ala Ala
420 425 430
Asn Leu Gln Asp Pro Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg
435 440 445
Arg Glu Glu Tyr Asp Val Leu Glu Lys Lys Arg Ala Arg Asp Pro Glu
450 455 460
Met Gly Gly Lys Gln Gln Arg Arg Arg Asn Pro Gln Glu Gly Val Tyr
465 470 475 480
Asn Ala Leu Gln Lys Asp Lys Met Pro Glu Ala Tyr Ser Glu Ile Gly
485 490 495
Thr Lys Gly Glu Arg Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln
500 505 510
Asp Ser His Phe Gln Ala Val Gln Phe Gly Asn Arg Arg Glu Arg Glu
515 520 525
Gly Ser Glu Leu Thr Arg Thr Leu Gly Leu Arg Ala Arg Pro Lys Gly
530 535 540
Glu Ser Thr Gln Gln Ser Ser Gln Ser Cys Ala Ser Val Phe Ser Ile
555 550 565 560
Pro Thr Leu Trp Ser Pro Trp Pro Pro Ser Ser Ser Ser Gln Leu
565 570 575

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 462 amino acids

(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Leu Val Leu Gln Leu
1 5 10 15
Ala Leu Leu Pro Ala Ala Thr Gln Gly Asn Lys Val Val Leu Gly Lys
20 25 30
Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln Lys Lys Ser
35 40 45
Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys Ile Leu Gly Asn
50 55 60
Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys Leu Asn Asp Arg Ala
65 70 75 80
Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe Pro Leu Ile Ile
85 90 95
Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr Tyr Ile Cys Glu Val Glu
100 105 110
Asp Gln Lys Glu Glu Val Gln Leu Leu Val Phe Gly Leu Thr Ala Asn
115 120 125
Ser Asp Thr His Leu Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu Glu
130 135 140
Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly
145 150 155 160
Lys Asn Ile Gln Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu
165 170 175
Gln Asp Ser Gly Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys
180 185 190
Val Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser
195 200 205
Ser Ile Val Tyr Lys Lys Glu Gly Glu Gln Val Glu Phe Ser Phe Pro
210 215 220
Leu Ala Phe Thr Val Glu Lys Leu Thr Gly Ser Gly Glu Leu Trp Trp
225 230 235 240
Gln Ala Glu Arg Ala Ser Ser Ser Lys Ser Trp Ile Thr Phe Asp Leu
245 250 255
Lys Asn Lys Glu Val Ser Val Lys Arg Val Thr Gln Asp Pro Lys Leu
260 265 270
Gln Met Gly Lys Lys Leu Pro Leu His Leu Thr Leu Pro Gln Ala Leu
275 280 285
Pro Gln Tyr Ala Gly Ser Gly Asn Leu Thr Leu Ala Leu Glu Ala Lys
290 295 300
Thr Gly Lys Leu His Gln Glu Val Asn Leu Val Val Met Arg Ala Thr
305 310 315 320
Gln Leu Gln Lys Asn Leu Thr Cys Glu Val Trp Gly Pro Thr Ser Pro
325 330 335
Lys Leu Met Leu Ser Leu Lys Leu Glu Asn Lys Glu Ala Lys Val Ser
340 345 350
Lys Arg Glu Lys Pro Val Trp Val Leu Asn Pro Glu Ala Gly Met Trp
355 360 365
Gln Cys Leu Leu Ser Asp Ser Gly Gln Val Leu Leu Glu Ser Asn Ile
370 375 380
Lys Val Leu Pro Thr Trp Ser Thr Pro Val His Ala Asp Pro Gln Leu
385 390 395 400
Cys Tyr Ile Leu Asp Ala Ile Leu Phe Leu Tyr Gly Ile Val Leu Thr
405 410 415
Leu Leu Tyr Cys Arg Leu Lys Ile Gln Val Arg Lys Ala Asp Ile Ala
420 425 430
Ser Arg Glu Lys Ser Asp Ala Val Tyr Thr Gly Leu Asn Thr Arg Asn

435 440 445
Gln Glu Thr Tyr Glu Thr Leu Lys His Glu Lys Pro Pro Gln
450 455 460 462

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 532 amino acids
(B) TYPE: amino acid
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Met Asn Arg Gly Val Pro Phe Arg His Leu Leu Val Leu Gln Leu
1 5 10 15
Ala Leu Leu Pro Ala Ala Thr Gln Gly Asn Lys Val Val Leu Gly Lys
20 25 30
Lys Gly Asp Thr Val Glu Leu Thr Cys Thr Ala Ser Gln Lys Lys Ser
35 40 45
Ile Gln Phe His Trp Lys Asn Ser Asn Gln Ile Lys Ile Leu Gly Asn
50 55 60
Gln Gly Ser Phe Leu Thr Lys Gly Pro Ser Lys Leu Asn Asp Arg Ala
65 70 75 80
Asp Ser Arg Arg Ser Leu Trp Asp Gln Gly Asn Phe Pro Leu Ile Ile
85 90 95
Lys Asn Leu Lys Ile Glu Asp Ser Asp Thr Tyr Ile Cys Glu Val Glu
100 105 110
Asp Gln Lys Glu Glu Val Gln Leu Leu Val Phe Gly Leu Thr Ala Asn
115 120 125
Ser Asp Thr His Leu Leu Gln Gly Gln Ser Leu Thr Leu Thr Leu Glu
130 135 140
Ser Pro Pro Gly Ser Ser Pro Ser Val Gln Cys Arg Ser Pro Arg Gly
145 150 155 160
Lys Asn Ile Gln Gly Gly Lys Thr Leu Ser Val Ser Gln Leu Glu Leu
165 170 175
Gln Asp Ser Gly Thr Trp Thr Cys Thr Val Leu Gln Asn Gln Lys Lys
180 185 190
Val Glu Phe Lys Ile Asp Ile Val Val Leu Ala Phe Gln Lys Ala Ser
195 200 205
Ser Ile Val Tyr Lys Lys Glu Gly Glu Gln Val Glu Phe Ser Phe Pro
210 215 220
Leu Ala Phe Thr Val Glu Lys Leu Thr Gly Ser Gly Glu Leu Trp Trp
225 230 235 240
Gln Ala Glu Arg Ala Ser Ser Ser Lys Ser Trp Ile Thr Phe Asp Leu
245 250 255
Lys Asn Lys Glu Val Ser Val Lys Arg Val Thr Gln Asp Pro Lys Leu
260 265 270
Gln Met Gly Lys Lys Leu Pro Leu His Leu Thr Leu Pro Gln Ala Leu
275 280 285
Pro Gln Tyr Ala Gly Ser Gly Asn Leu Thr Leu Ala Leu Glu Ala Lys
290 295 300
Thr Gly Lys Leu His Gln Glu Val Asn Leu Val Val Met Arg Ala Thr
305 310 315 320
Gln Leu Gln Lys Asn Leu Thr Cys Glu Val Trp Gly Pro Thr Ser Pro
325 330 335
Lys Leu Met Leu Ser Leu Lys Leu Glu Asn Lys Glu Ala Lys Val Ser
340 345 350
Lys Arg Glu Lys Pro Val Trp Val Leu Asn Pro Glu Ala Gly Met Trp
355 360 365

Gln Cys Leu Leu Ser Asp Ser Gly Gln Val Leu Leu Glu Ser Asn Ile
370 375 380
Lys Val Leu Pro Thr Trp Ser Thr Pro Val His Ala Asp Pro Lys Leu
385 390 395 400
Cys Tyr Leu Leu Asp Gly Ile Leu Phe Ile Tyr Gly Val Ile Leu Thr
405 410 415
Ala Leu Phe Leu Arg Val Lys Phe Ser Arg Ser Ala Glu Pro Pro Ala
420 425 430
Tyr Gln Gln Gly Gln Asn Gln Leu Tyr Asn Glu Leu Asn Leu Gly Arg
435 440 445
Arg Glu Glu Tyr Asp Val Leu Asp Lys Arg Arg Gly Arg Asp Pro Glu
450 455 460
Met Gly Gly Lys Pro Arg Arg Lys Asn Pro Gln Glu Gly Leu Tyr Asn
465 470 475 480
Glu Leu Gln Lys Asp Lys Met Ala Glu Ala Tyr Ser Glu Ile Gly Met
485 490 495
Lys Gly Glu Arg Arg Gly Lys Gly His Asp Gly Leu Tyr Gln Gly
500 505 510
Leu Ser Thr Ala Thr Lys Asp Thr Tyr Asp Ala Leu His Met Gln Ala
515 520 525
Leu Pro Pro Arg
530

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

CGCGGGGTGA CGGTGCCCTC CAGCAGCTTG GGC

33

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 50 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

CGCGGGGATC CGTCGTCCAG AGCCCGTCCA GCTCCCCGTC CTGGGCCTCA

50

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

CGCGGGCGGC CGCGACGCCG GCCAAGACAG CAC

33

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

CGCGTTGACG AGCAGCCAGT TGGGCAGCAG CAG

33

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 15 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

CGCGGGCGGC CGCTA

15

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

CGCGGGCTCG TTATAGAGCT GGTTCTGGCG CTGCTTCTTC TG

42

(2) INFORMATION FOR SEQ ID NO:13:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 48 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

CGCGGGGAGC TCGTTATAGA GCTGGTTTGC CGCCGAATTC TTATCCCG

48

(2) INFORMATION FOR SEQ ID NO:14:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

CGCGGGGCGG CCACCGGTCC TCGCCAGCAC ACA

33

(2) INFORMATION FOR SEQ ID NO:15:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 36 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

CGCGGGACGC GTTCAGCCG TCCTCGCCAG CACACA

36

(2) INFORMATION FOR SEQ ID NO:16:

- (i) SEQUENCE CHARACTERISTICS:
- (A) LENGTH: 33 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

CGCGGGACGC GTGACCCCTGA GATGGGGGAA AAG

33

(2) INFORMATION FOR SEQ ID NO:17:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 33 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CGCGGGACGC GTATTGGGAT GAAAGGCGAG CGC

33

(2) INFORMATION FOR SEQ ID NO:18:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 26 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

CCCGGATCCC AGCATGGGCA GCTCTT

26

(2) INFORMATION FOR SEQ ID NO:19:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 42 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

CGCGGGGCGG CCGCTTTAGT TATTACTGTT GACATGGTCG TT

42

(2) INFORMATION FOR SEQ ID NO:20:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 30 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

GCGGGGGGAT CCCACTGTCC AAGCTCCCAG

30

(2) INFORMATION FOR SEQ ID NO:21:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 32 bases
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:21;

GCGGGGGCGG CCGCCTAAAT ACGGTTCTGG TC

32

(2) INFORMATION FOR SEQ ID NO:22:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

TCAGAAAGAG ACAACCTGAA GAAACCAACA λ

31

(2) INFORMATION FOR SEQ ID NO:23:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 31 bases
- (B) TYPE: nucleic acid
- (C) STRANDEDNESS: single
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: nucleic acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

TTGTTGGTTT CTTCAGGTTG TGTCTTCCTG A

31

(2) INFORMATION FOR SEQ ID NO:24:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 171 amino acids
- (B) TYPE: amino acid
- (C) TOPOLOGY: linear

(ii) MOLECULE TYPE: amino acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

Met Glu His Ser Thr Phe Leu Ser Gly Leu Val Leu Ala Thr Leu Leu
5 10 15
Ser Gln Val Ser Pro Phe Lys Ile Pro Ile Glu Glu Leu Glu Asp Arg
20 25 30
Val Phe Val Asn Cys Asn Thr Ser Ile Thr Trp Val Glu Gly Thr Val
35 40 45
Gly Thr Leu Leu Ser Asp Ile Thr Arg Leu Asp Leu Gly Lys Arg Ile
50 55 60
Leu Asp Pro Arg Gly Ile Tyr Arg Cys Asn Gly Thr Asp Ile Tyr Lys
65 70 75 80
Asp Lys Glu Ser Thr Val Gln Val His Tyr Arg Met Cys Gln Ser Cys
85 90 95
Val Glu Leu Asp Pro Ala Thr Val Ala Gly Ile Ile Val Thr Asp Val
100 105 110
Ala Ile Thr Leu Leu Ala Leu Gly Val Phe Cys Phe Ala Gly His
115 120 125
Glu Thr Gly Arg Leu Ser Gly Ala Ala Asp Thr Gln Ala Leu Leu Arg
130 135 140
Asn Asp Gln Val Tyr Gln Pro Leu Arg Asp Arg Asp Asp Ala Gln Tyr
145 150 155 160
Ser His Leu Gly Gly Asn Trp Ala Arg Asn Lys
165 170

(2) INFORMATION FOR SEQ ID NO:25:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 182 amino acids
- (B) TYPE: amino acid
- (C) TOPOLOGY: linear

(ii) MOLECULE TYPE: amino acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

Met Glu Gln Gly Lys Gly Leu Ala Val Leu Ile Leu Ala Ile Ile Leu
5 10 15
Leu Gln Gly Thr Leu Ala Gln Ser Ile Lys Gly Asn His Leu Val Lys
20 25 30
Val Tyr Asp Tyr Gln Glu Asp Gly Ser Val Leu Leu Thr Cys Asp Ala
35 40 45
Glu Ala Lys Asn Ile Thr Trp Phe Lys Asp Gly Lys Met Ile Gly Phe
50 55 60

Leu Thr Glu Asp Lys Lys Trp Asn Leu Gly Ser Asn Ala Lys Asp
65 70 75 80
Pro Arg Gly Met Tyr Gln Cys Lys Gly Ser Gln Asn Lys Ser Lys Pro
85 90 95
Leu Gln Val Tyr Tyr Arg Met Cys Gln Asn Cys Ile Glu Leu Asn Ala
100 105 110
Ala Thr Ile Ser Gly Phe Leu Phe Ala Glu Ile Val Ser Ile Phe Val
115 120 125
Leu Ala Val Gly Val Tyr Phe Ile Ala Gly Gln Asp Gly Val Arg Gln
130 135 140
Ser Arg Ala Ser Asp Lys Gln Thr Leu Leu Pro Asn Asp Gln Leu Tyr
145 150 155 160
Gln Pro Leu Lys Asp Arg Glu Asp Asp Gln Tyr Ser His Leu Gln Gly
165 170 175
Asn Gln Leu Arg Arg Asn
180

(2) INFORMATION FOR SEQ ID NO:26:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 220 amino acids
- (B) TYPE: amino acid
- (C) TOPOLOGY: linear

(ii) MOLECULE TYPE: amino acids

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

Met Pro Gly Gly Leu Glu Ala Leu Arg Ala Leu Pro Leu Leu Phe
5 10 15
Leu Ser Tyr Ala Cys Leu Gly Pro Gly Cys Gln Ala Leu Arg Val Glu
20 25 30
Gly Gly Pro Pro Ser Leu Thr Val Asn Leu Gly Glu Glu Ala Arg Leu
35 40 45
Thr Cys Glu Asn Asn Gly Arg Asn Pro Asn Ile Thr Trp Trp Phe Ser
50 55 60
Leu Gln Ser Asn Ile Thr Trp Pro Pro Val Pro Leu Gly Pro Gly Gln
65 70 75 80
Gly Thr Thr Gly Gln Leu Phe Phe Pro Glu Val Asn Lys Asn Thr Gly
85 90 95
Ala Cys Thr Gly Cys Gln Val Ile Glu Asn Asn Ile Leu Lys Arg Ser
100 105 110
Cys Gly Thr Tyr Leu Arg Val Arg Asn Pro Val Pro Arg Pro Phe Leu
115 120 125
Asp Met Gly Glu Gly Thr Lys Asn Arg Ile Ile Thr Ala Glu Gly Ile
130 135 140
Ile Leu Leu Phe Cys Ala Val Val Pro Gly Thr Leu Leu Leu Phe Arg
145 150 155 160
Lys Arg Trp Gln Asn Glu Lys Phe Gly Val Asp Met Pro Asp Asp Tyr
165 170 175
Glu Asp Glu Asn Leu Tyr Glu Gly Leu Asn Leu Asp Asp Cys Ser Met
180 185 190
Tyr Glu Asp Ile Ser Arg Gly Leu Gln Gly Thr Tyr Gln Asp Val Gly
195 200 205
Asn Leu His Ile Gly Asp Ala Gln Leu Glu Lys Pro
210 215 220

(2) INFORMATION FOR SEQ ID NO:27:

(i) SEQUENCE CHARACTERISTICS:

- (A) LENGTH: 228 amino acids
- (B) TYPE: amino acid
- (D) TOPOLOGY: linear

(ii) MOLECULE TYPE: amino acid

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

Met Ala Thr Leu Val Ser Ser Met Pro Cys His Trp Leu Leu Phe
5 10 15
Leu Leu Leu Phe Ser Gly Glu Pro Val Pro Ala Met Thr Ser Ser
20 25 30
Asp Leu Pro Leu Asn Phe Gln Gly Ser Pro Cys Ser Gln Ile Trp Gln
35 40 45
His Pro Arg Phe Ala Ala Lys Lys Arg Ser Ser Met Val Lys Phe His
50 55 60
Cys Tyr Thr Asn His Ser Gly Ala Leu Thr Trp Phe Arg Lys Arg Gly
65 70 75 80
Ser Gln Gln Pro Gln Glu Leu Val Ser Glu Glu Gly Arg Ile Val Gln
85 90 95
Thr Gln Asn Gly Ser Val Tyr Thr Leu Thr Ile Gln Asn Ile Gln Tyr
100 105 110
Glu Asp Asn Gly Ile Tyr Phe Cys Lys Gln Lys Cys Asp Ser Ala Asn
115 120 125
His Asn Val Thr Asp Ser Cys Gly Thr Glu Leu Leu Val Leu Gly Phe
130 135 140
Ser Thr Leu Asp Gln Leu Lys Arg Arg Asn Thr Leu Lys Asp Gly Ile
145 150 155 160
Ile Leu Ile Gln Thr Leu Leu Ile Ile Leu Phe Ile Ile Val Pro Ile
165 170 175
Phe Leu Leu Leu Asp Lys Asp Asp Gly Lys Ala Gly Met Glu Glu Asp
180 185 190
His Thr Tyr Glu Gly Leu Asn Ile Asp Gln Thr Ala Thr Tyr Glu Asp
195 200 205
Ile Val Thr Leu Arg Thr Gly Glu Val Lys Trp Ser Val Gly Glu His
210 215 220
Pro Gly Gln Glu
225